

"Relations all the way" means what?



What are Tables?

Tables are just a form of Structured Data Representation, popularized by Spreadsheets (as Grids) and SQL-Compliant Database Management Systems (DBMS) where Entity Relationships Types are represented as Records in a Table.

Each Table Name identifies a Relational Set (Relation) operated on by a given SQL-Compliant DBMS.



What is Comma Separated Values (CSV)?

A Structured Data Interchange format (created by IBM) for Data Export and Import across Database Management Systems (DBMS) that operate on Entity Relationships represented as Records in a Table.

A majority of modern SQL-compliant DBMS products support this interchange format.



What are RDF Graphs?

An RDF Sentence is a 3-part structure comprising a Subject, Predicate, and Object that represents Entity Relationship Types grouped by Sentence Predicate.

Each Predicate identifies a Relational Set (Relation) operated on by a given RDF-Compliant DBMS (or Store).



What are RDF Notations & Document Types?

As an Abstract Data Definition Language, RDF is expressible using a variety of notations. Examples include:

- RDF-Turtle Notation
- RDF-JSON Notation
- JSON-LD Notation
- others

Data expressed using various RDF Notations can be transformed and persisted to a number of Document Content Types. Examples include:

- text/turtle
- application/json
- application/ld+json





What is Relational Table to RDF Graph Transformation?

How Entity Relationship Types represented as Records in a Table are transformed into Entity Relationship Types represented as RDF-based Subject, Predicate, Object sentences grouped by Predicate.

For Example:

CSV to RDF Document transformation results in RDF-based Entity Relationship Graphs that manifests as the following document content-types: "text/turtle" (RDF-Turtle), "application/json" (RDF-JSON), and "application/json+ld" (JSON-LD), etc..



CSV to RDF Transformation Exercise



Demonstration, Step 1 — Open Document of Interest

Furthermore, each entity class is accompanied by three sets of statistics files (downloadable as .zip below): Summary Statistics, Column Distribution and Table Statistics. Each of these files contains statistics about different dimensions of the table set of the specific class:

- The Summary Statistics files contain summary statistics about the tables set, such as overall number of tables, overall number of rows, median rows per table and also information about row distribution in the form of quantile borders.
- The Column Distribution files contain an overview over the columns and their percentual occurrence across all tables of that entity class, similar to Table 5.
- The Table Statistics files contain a list of all tables. For each table, the file reports the number of rows, the number of columns as well as the density of each column. The table statistics file allows you to identify all tables of a class containing a specific column, e.g. all tables containing a GTIN column, or all tables containing a phone number column having a minimum density of 80%.

Class	Top 100	Minimum3	Rest	JSON sample	CSV sample	Statistics Files
Product	top100	minimum3	rest	sample json	sample_csv	statistics
Person	top100	minimum3	rest	sample json	sample csv	statistics
LocalBusiness	top100	minimum3	rest	sample json	sample csv	statistics
CreativeWork	top100	minimum3	rest	sample json	sample csv	statistics
Event	top100	minimum3	rest	sample json	sample csv	statistics
Place	top100	minimum3	rest	sample json	sample csv	statistics
Restaurant	top100	minimum3	rest	sample json	sample csv	statistics
Recipe	top100	minimum3	rest	sample json	sample csv	statistics
JobPosting	top100	minimum3	rest	sample json	sample csv	statistics
Hotel	top100	minimum3	rest	sample json	sample csv	statistics
MusicRecording	top100	minimum3	rest	sample json	sample csv	statistics
Book	top100	minimum3	rest	sample json	sample csv	statistics
MusicAlbum	top100	minimum3	rest	sample json	sample csv	statistics
Movie	top100	minimum3	rest	sample json	sample csv	statistics
GeoCoordinates	top100	minimum3	rest	sample json	sample csv	statistics



Demonstration, Step 2 — Discover, Transform, and Present Knowledge Graph

CSV	
Statement Collection #9	
Entity	_:b0
Attributes	
#row_id	O(xsd:integer)
#name	Goldline Coloured Card Whte.Gcc3220 Pk25
#offers	{"category": "Paper", "sku": "329235", "description": "The Goldine Plain White Card uses heavy 210gsm card s
	and artwork. It is also suitable for copiers and laser printers to create large format printouts. The card comes in a for eye-catching displays and posters.", "pricecurrency": "GBP"}
#description	null
<u>#sku</u>	null
<u>#url</u>	null
#brand	{"name": "Goldline"}
<u>#page_url</u>	https://webstore.astute.uk.com/goldline-coloured-card-whte-gcc3220-pk25/pi/329235-110250-1.html (xsd:an
Statement Collection #10	
Entity	_:b1
Attributes	
#row_id	1(<u>xsd:integer</u>)
#name	UNCONDITIONAL long shirt with centre back tail split in silver grey light tencel shirting
<u>#offers</u>	{"description": "UNCONDITIONAL SS16 long shirt with centre back tail split in silver grey light tencel shirting, fe black buttons * Also available in black Colour : silver grey Product code : SH605 Care instructions : Dry clean red medium, tumble dry medium, do not bleach.", "price": "£169", "pricecurrency": "GBP"}
#description	null
#sku	null
#url	https://www.unconditional.uk.com/products/unconditional-long-shirt-with-centre-back-tail-split-in-silver-grey-l
#brand	null
#page_url	https://www.unconditional.uk.com/products/unconditional-long-shirt-with-centre-back-tail-split-in-silver-grey-l
Statement Collection #11	
Entity	_:b2
Attributes	02
#row_id	2(xsd:integer)
#name	Q-Connect Epson T1291 Remanufactured Black Inkjet Cartridge C13T12914011 / T129140
#offers	{"category": "Printer Supplies#Cartridges", "sku": "OB12914", "description": "With all the quality of an Epson ca
	Epson Stylus Black Inkjet Cartridge offers a standard of excellence you would expect from a branded product wi flawlessly with your printer, this cartridge contains professional-grade ink that creates vivid and consistently stu cartridge.", "pricecurrency": "GBP"}
#description	null
#sku	null
#url	null
#brand	{"name": "Q Connect"}
#page url	https://webstore.astute.uk.com/g-connect-epson-t1291-remanufactured-black-inkiet-cartridge-c13t12914011
OpenLink Structured Data Sniffer	

d stock that is ideal for sturdy, long-lasting posters in a size between A1 and A2 - 500 x 700 mm - great

anyURI)

features a concealed button placket and contrast recommended, wash separately in warm water, iron

/-light-tencel-shirting

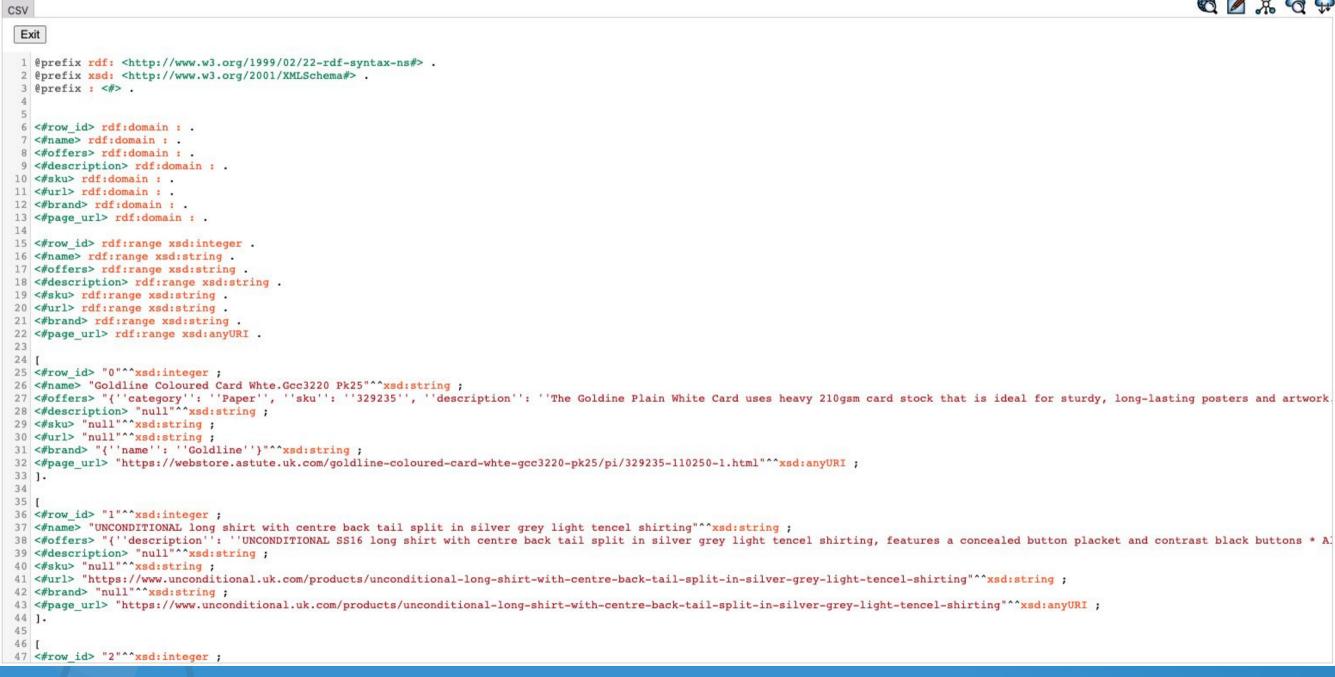
y-light-tencel-shirting (xsd:anyURI)

n cartridge as a fraction of the price, the Q-Connect without the hefty price tag. Designed to work stunning prints for the entire life cycle of the

1---t129140/pi/OB12914-1612328-1.html (Copyright © 2015-2021 <u>OpenLink Software</u>



Demonstration, Step 3 — View Transformation to a **Specific RDF Notation, such as RDF-Turtle**



0 × 0 ÷



Demonstration, Step 4 — Query Knowledge Graph using SPARQL

With OSDS, you are only a single click away from generating a SPARQL query results page from any CSV document you've opened in your browser.

Click on any Hyperlink for deeper exploration of the underlying Knowledge Graph!

EntityID	count	EntityTypeID
https://linkeddata.uriburner.com/proxy-iri/35751b9cdd64b60f18eecbd4fdeee897b794cbdc	24	http://www.openlink
https://linkeddata.uriburner.com/proxy-iri/35751b9cdd64b60f18eecbd4fdeee897b794cbdc	24	http://data.dws.infor mannheim.de/struc
http://www.w3.org/2002/07/owl#	8	http://www.openlink
http://data.dws.informatik.uni- mannheim.de/structureddata/schemaorgtables/Product/samples/Product_csvsample.csv#class	1	http://www.w3.org/2
http://data.dws.informatik.uni- mannheim.de/structureddata/schemaorgtables/Product/samples/Product_csvsample.csv		http://schema.org/C
https://linkeddata.uriburner.com/about/id/entity/http/data.dws.informatik.uni- mannheim.de/structureddata/schemaorgtables/Product/samples/Product_csvsample.csv	1	http://www.openlink

ksw.com/schemas/opltable#Row

ormatik.unictureddata/schemaorgtables/Product/samples/Product_csvsample.csv#class hksw.com/schema/attribution#DataSource

/2000/01/rdf-schema#Class

CreativeWork

ksw.com/schemas/opltable#Table



Demonstration, Step 5 — View Query Editor for OSDS-Generated SPARQL Query

← → C ☆ 🔒 https://linkeddata.uriburner.com/sparql/?query=DEFINE%20get%3Asot	ft%20"soft"%20% 🎛 🚖 🌍 🗹 🌒 🔻 🎴 🗐 📚 💁 🕤
👖 Apps 🔇 Verizon Router 🔥 Windows Azure 🔇 RoosterJs Demo P 🔇 Latest License Ge	<pre>① 1 DEFINE get:soft "soft" 2 SELECT (SAMPLE(2s) AS 2EntityID)</pre>
	<pre>① 1 DEFINE get:soft "soft" 2 SELECT (SAMPLE(?s) AS ?EntityID) 3 (COUNT(*) AS ?count) 4 (?o AS ?EntityTypeID) 5 FROM <http: data.dws.informatik.uni-mannheim.de="" structuredda<br="">6 WHERE { 7 ?s a ?o . 8 FILTER (CONTAINS(STR(?o), 'schema')) 9 } 10 GROUP BY ?o 11 ORDER BY DESC (?count) 12 LIMIT 50 13</http:></pre>





What's Actually Going on Here?

Virtuoso, our Multi-Model Relational Engine, handles Relations represented as Tables and/or RDF Graphs. Simply import or attach any CSV, ODBC, or JDBC data source, and Virtuoso does the rest!

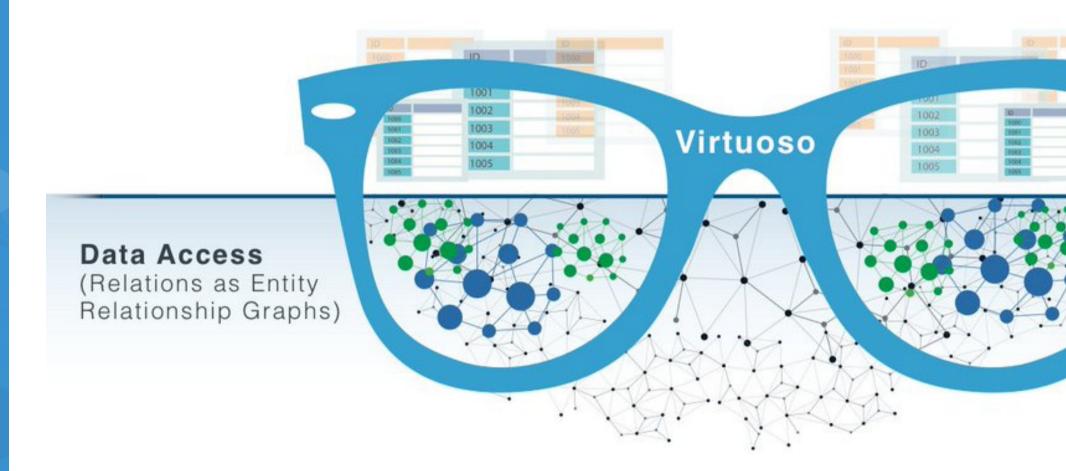
Fundamentally —

- <u>Tables</u> are powerful visualization tools for data analytics
- <u>Graphs</u> are powerful navigational tools for query construction and optimization

All this powerful functionality is packed into *one-click interactions* initiated from OSDS.



Virtuoso HyperData "Context Lenses"



Data Presentation

(Relations as Tables)



How Do I Get Going?

• OSDS Browser Extension - various options for installation

• Virtuoso On Premise — Installer and Free 30-day Evaluation Licenses from our <u>Download Site</u>

Virtuoso Docker Container — various containers from <u>Docker Hub</u>

• Virtuoso Cloud — Bring-Your-Own License (BYOL) and Pay-As-You-Go (PAGO) instances in the Cloud at Amazon Web Services (AWS) and Microsoft Azure



ADDITIONAL INFORMATION

- <u>OpenLink Structured Data Sniffer</u>
- <u>Virtuoso Home Page</u>
- <u>Virtuoso Weblog</u>
- <u>Virtuoso Community Forum</u>

